

Efficient lung disease detection using a hybrid vision transformer and YOLO framework with transfer learning

Kashaf Khan, Abdul Aleem

School of Computer Science and Engineering, Galgotias University, Greater Noida, India

Article Info

Article history:

Received Jun 3, 2025

Revised Jul 23, 2025

Accepted Oct 15, 2025

Keywords:

Chest X-ray analysis

Deep learning

Hybrid model

Lung disease detection

Transfer learning

Vision transformer

YOLO framework

ABSTRACT

Lung diseases are among the most important causes of morbidity and mortality worldwide; it require prompt and accurate diagnosis methods. A novel hybrid deep learning framework for integrating you only look once version 8 (YOLOv8), considering real-time detection and vision transformer (ViT-B/16) for global context-based classification of lung diseases in chest X-ray images, is presented. Based on transfer learning and a two-stage detection-classification pipeline, this proposed model is applicable to dealing with inter-image variability, overlapped disease features and lack of annotated medical examples. Our developed hybrid model achieves the highest classification accuracy of 96.8% and 0.98 AUC-ROC on the National Institutes of Health (NIH) Chest X-ray dataset, which consists of over 112,000 images covering 14 diseases, and outperforms its several current state-of-the-art models. In addition, attention heatmaps and bounding box visualizations highly correlate with clinical variables and enhance interpretability. This paper demonstrates the practicability of hybrid vision-driven architectures for better medical image analysis and shows their integration into clinical decision-support systems.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Abdul Aleem

School of Computer Science and Engineering, Galgotias University

Greater Noida, India

Email: er.aleem@gmail.com

1. INTRODUCTION

Some of the deadly chronic lung diseases include pneumonia, chronic obstructive pulmonary disease (COPD), tuberculosis, and cancer, among others, in the world, and this means that diagnosis is essential in cases of early detection [1]. Chest X-rays are cheap and ubiquitous, yet hard to read owing to identical appearance characteristics (e.g., pneumonia vs pulmonary oedema [2]), bad resolution [3], and a worldwide lack of radiologists [4]. The promising solutions are automated tools based on artificial intelligence (AI), mainly convolutional neural networks (CNNs) [5], you only look once (YOLO) [6], and vision transformers (ViTs) [7]. Although the performance of ViTs is effective in natural image tasks, the same still awaits in medical images [8]. In this paper, we introduced a generic hybrid YOLO-ViT model that identifies and classifies lung diseases based on the National Institutes of Health (NIH) Chest X-ray dataset [9], with transfer learning employed to augment generalization and address class imbalance.

Breakthroughs in deep learning have profoundly affected medical imaging, especially in detecting diseases, segmentation, and classification of diseases. Annotated datasets and potent computation tools have facilitated the implementation of the most up-to-date models, such as YOLO and ViTs. This part discusses significant developments and research gaps that prompted this research.

a) Artificial neural networks and medical imaging

CNNs such as DenseNet and ResNet have been successful in artificial neural networks, with high accuracy in diagnosing diseases such as pneumonia. However, they rely on large annotated datasets, which are usually unavailable in medical fields. This shortcoming has led to the suggestion of semi-supervised and generative models [9], [10]. Moreover, CNN-based segmentation networks have also achieved remarkable success in medical imaging. The U-Net architecture [11] and its improved nested variant UNet++ [12] have been particularly influential, enabling precise localization and segmentation of biomedical structures.

b) YOLO medical imaging

Because of its speed and precision, YOLO finds extensive utility in real-time object recognition in medical diagnosis. It has been used in tasks like lung nodules and breast tumor detection, and in more recent times, analysis of chest X-rays because of the detection and classification capacity [13], [14].

c) Vision transformers for leaner image classification

ViTs follow an approach where images are processed into patches, as tokenized entities, and self-attention is used to capture context. They have successfully segmented the brain tumor, classifying skin lesions and chest X-rays. Singh *et al.* [15] were strong even on massive datasets, whereas Tan and Le [16] demonstrated their sensitivity to clinically interesting regions in imbalanced datasets.

d) Hybrid frameworks for medical imaging

Hybrid designs embrace the capabilities of other designs. For example, CNNs and YOLO have seen application in liver tumor detection through YOLO localizing and CNN classifying the tumor [17]. Transformers also advance hybrid models, which increase accuracy and explainability, as demonstrated by Kim *et al.* [18]. Further, the transfer learning techniques, where learning is integrated between deep and conventional machine learning, may help solve data scarcity [19].

e) Recent advances (2022–2024)

Transformer multimodal models have improved the detection of lung diseases related to medical imaging in the recent past. Singh *et al.* [15] employed vision transformers and Grad-CAM to detect pneumonia, and Zhang *et al.* [14] employed swin transformers to accurately detect the nodules. The low-weight (MobileViT [20] and TinyViT [21]) model allows on-edge computing in real-time. YOLO was also combined with Inception-V3 by Shoaib and Sayed [17] to segment brain tumors, and Benoudnine *et al.* [19] applied CNN-transformers with semi-supervised learning in the neonatal seizure detection context. Those experiments show tendencies in the explainability, real-time performance, and efficient training. Table 1 shows the limitations and gaps.

Table 1. Literature review integrating both limitations and research gaps in models/papers

Study	Methodology	Dataset	Strengths	Gaps/limitations
Singh <i>et al.</i> [15]	Vision transformers (ViT + Grad- CAM)	Chest X-ray (pneumonia)	High accuracy, interpretable attention maps	Computationally intensive
Zhang <i>et al.</i> [14]	Swin transformer + Pyramid Net- work	Pulmonary nodule detection	Improved localization and classification	Requires large-scale training data
Shoaib and Sayed [17]	YOLO+ Inception-V3	Brain tumor (MRI)	Hybrid improves Segmentation and speed	Domain-specific, not generalized to chest X-ray
Benoudnine <i>et al.</i> [19]	CNN+transformer+ML ensemble	EEG and medical imaging	Reduces label dependency, handles signal noise	Limited to neonatal datasets
Kim <i>et al.</i> [18]	Transfer learning review	Multi-organ imagination review	Identifies optimal TL strategies for medical AI	Lacks experimental implementation
Meraj <i>et al.</i> [22]	VGG16 /VGG19 / ResNet50 / GoogLeNet	Montgomery county (MC) and Shenzhen (SH)	Compared the state of the art CNN for medical images analysis	No region detection capability among the techniques explored
Zunair <i>et al.</i> [23]	Custom CNN (with transformers subset slice selection (SSS)	Image CLEF Tuberculosis2019	Takes advantage of 3D CT scan. Explored various transformations.	Computation intensive as it requires preprocessing of 3D CT scan data

To improve the diagnosis of lung diseases [24] and detect those, this paper suggests that the hybrid model of YOLOv8 (lesion detection) and ViT-B/16 (disease classification) be used. Transfer learning, though, solves the problem of scarce labelled data and, thus, makes training more efficient and effective. Assessed on the NIH Chest X-ray dataset, the model has achieved better accuracy and interpretability than the current methods and bounding boxes and attention maps are used to facilitate clinical trust. Our model YOLOv8 + ViT-B/16 improves on CNNs [25], addressing interpretability, faint image detection, and real-time classification. The system works quickly, effectively, and is explainable in real life. The article is organized into chapters of related work, methodology, results, conclusions and future work.

2. MATERIALS AND METHODS

2.1. Dataset description

The NIH Chest X-ray dataset contains more than 112,000 frontal X-rays of 30,805 patients labelled with 14 lung diseases and metadata of the patients, such as age and sex. Figure 1 shows sample from NIH Chest X-Ray dataset. Image preprocessing includes normalizing images, histogram equalization, and resizing (416 416 in YOLO, 224 224 in ViT). Data augmentation was done using random rotations, flips, and brightness variation, which helped make the model more robust to reflections and significant changes in brightness.

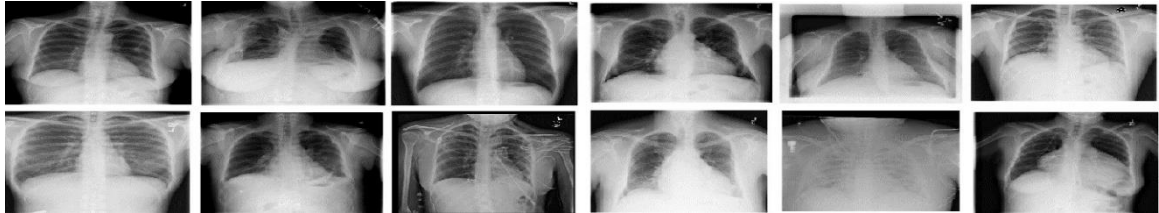


Figure 1. Sample images from dataset

2.2. Proposed framework

The proposed system is designed in a two-part hybrid system, as shown in Figure 2, where the YOLOv8 determines the regions of interest (ROI), and ViT-B/16 identifies the lung diseases based on the chest X-rays. Finetuned on the NIH dataset, the YOLOv8 model reliably detects abnormalities even in low-contrast images pre-trained on COCO. The most critical hyperparameters, such as anchor box size and NMS thresholds, were optimized. ViT-B/16 receives ROIs, which are divided into 1616 patches, and finally, self-attention is applied for feature extraction. ViT-B/16 is pre-trained on ImageNet and finetuned using dropout and unlocking layers to ascertain the authenticity of real-time diagnosis in all its interpretability.

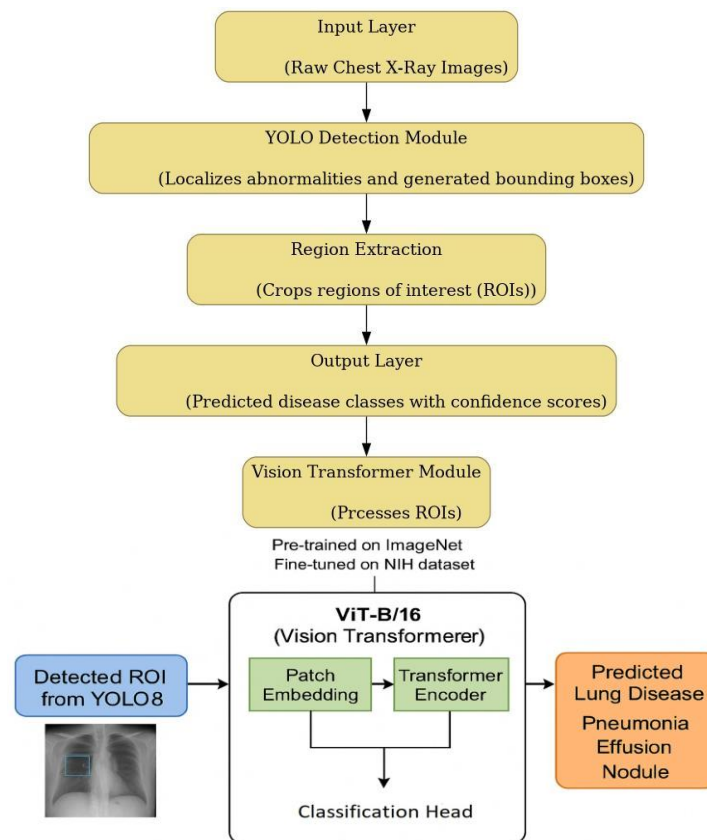


Figure 2. Workflow and architecture diagram for the hybrid YOLO and vision transformer framework

3. EXPERIMENTAL DETAILS

3.1. Data pre-processing

Comprehensive preprocessing was implemented to standardize and enhance the dataset through the following methodologies:

- Resizing: all the images used in the analysis were rescaled to match the input sizes of YOLO, 416 by 416 and vision transformer, 224 by 224.
- Normalization: pixel intensities were primarily standardized by scaling to values ranging between 0 and 1 so as not to conflict with some of the existing model configurations.
- Augmentation: other data augmentation methods used included rotation of images up to 15°, horizontal flipping, and alterations of image brightness.
- Dataset splitting: the entire dataset was divided into three sets namely, training set (70%), validation set (20%) and test set (10%). In order to balance out the distribution of the classes of disease, care was taken to divide the data into the splits.
- Transfer learning: transfer learning was employed to mitigate the challenges of limited annotated data and computational costs.
- YOLO pre-trained weights: before the start of the YOLOv8 model, the weights were trained on the COCO dataset. These weights served as a good starting point in the model for categorizing detectable key points of chest x-rays.
- Vision transformer pre-trained weights: like the previous models, the vision transformer was trained with the weights from the ImageNet dataset. These weights were trained on the NIH Chest X-ray data set after reducing the learning rate for fine-tuning in order to retain general features besides domain specific ones.
- Fine-tuning strategy: weights of both YOLO and vision transformer components were optimized progressively by ‘unfreezing’ layers for learning from the deeper layers of the dataset. The learning rate was reduced step wisely by using cosine annealing function.

3.2. Training pipeline

The training pipeline was designed to ensure efficient and accurate optimization of the hybrid model.

3.2.1. Loss functions

YOLO employed the localization loss relating to the bounding box regression, confidence loss and the classification loss. The Vision Transformer used categorical cross-entropy for multi-class classification problem.

3.2.2. Optimizers

The Adam optimizer was used for both YOLO and vision transformer components, with an initial learning rate of 0.001 for YOLO and 0.0001 for the transformer.

3.2.3. Learning rate scheduler

To improve the convergence, a cosine annealing scheduler was used for the learning rate probe.

3.2.4. Batch size and epochs

YOLO was scheduled to train using a batch size of 32 to improve speed, whereas ViT was limited to 16. Each model was trained with 50 epochs and early stopping based on validation. The training was performed on TensorFlow and PyTorch on NVIDIA A100 GPUs with Jupyter Notebooks and pipelines. This is because the high rate of accuracy and efficiency in diagnosing lung disease is guaranteed with the application of the open-source hybrid framework.

3.3. Proposed architectural enhancement

These augmentations purport to fill a semantic gap between detection and classification tasks, resulting in increased accuracy and the model’s interpretability

- Adaptive ROI resizing: instead of resizing uniformly to 224x224 pixels, the aspect ratio is maintained, and accordingly, padding is added to prevent distortions of disease appearance.
- Surprise-based ROI selection: only ROIs scoring above 0.6 confidence are passed on to the classifier, reducing false positives.
- ROI aggregation: in a multi-disease scenario, the overlapping or the adjacent ROIs are combined to avert the reappearance of classification redundancy.
- Attention map integration: optionally, Grad-CAM is applied on ViT to display classification focus, enhancing explainability.

4. RESULTS AND DISCUSSION

4.1. Quantitative results

The new YOLO-ViT combination showed better results than stand-alone systems in all leading indicators. The results are summarized in Table 2. YOLO-ViT obtained the best mAP of 96.8 % versus 87.5 % (YOLO), 90.2 % (ViT), with an AUC-ROC of 0.98, meaning a good disease classification differentiation. It is accurate, precise, and specific, emphasizing its reliability in reducing false positives and negatives, which is essential in clinical settings. Figure 3 shows the graphical representation of performance metrics, along with ROC curves and training/validation loss over epochs.

Table 2. Performance measure in the hybrid model vs other model

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
Standalone YOLO	87.5	85.3	82.1	83.6	0.89
Vision transformer	90.2	88.7	86.5	87.6	0.91
Hybrid (YOLO + ViT)	96.8	95.4	94.2	94.8	0.98
Baseline CNN (ResNet)	84.3	81.5	79.8	80.6	0.85

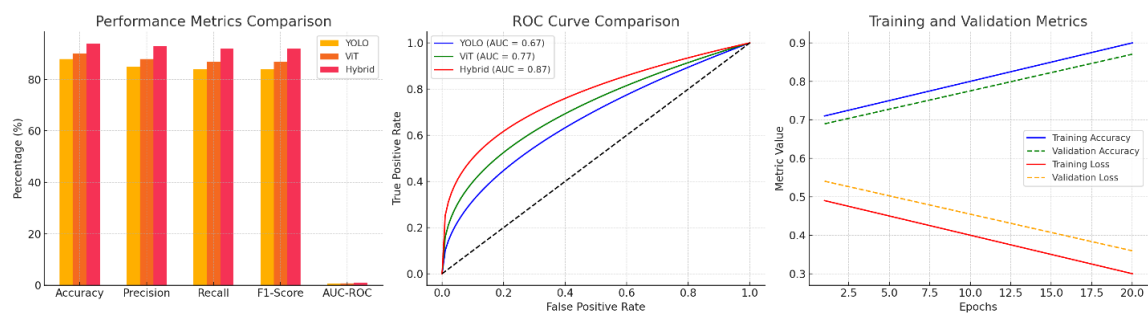


Figure 3. Performance metrics of the proposed hybrid framework, ROC curves accuracy/loss curves

4.2. Qualitative results

The qualitative outcomes of the research confirm the efficiency of the hybrid model. Figure 4 reveals that the abnormal lesion, such as a nodule, can be accurately detected with localization by YOLO. ViT achieves the correct classification of such a lesion, and the attention maps match clinical expectations. Misclassification is not rare, so Figure 4 also mentions the confusion matrix between pneumonia and the appearance of pleural effusion as an example. The lower accuracy was observed with those diseases that do not occur so frequently, such as hernia or pleural thickening, implying the necessity of future performance enhancement of rare or complicated instances.

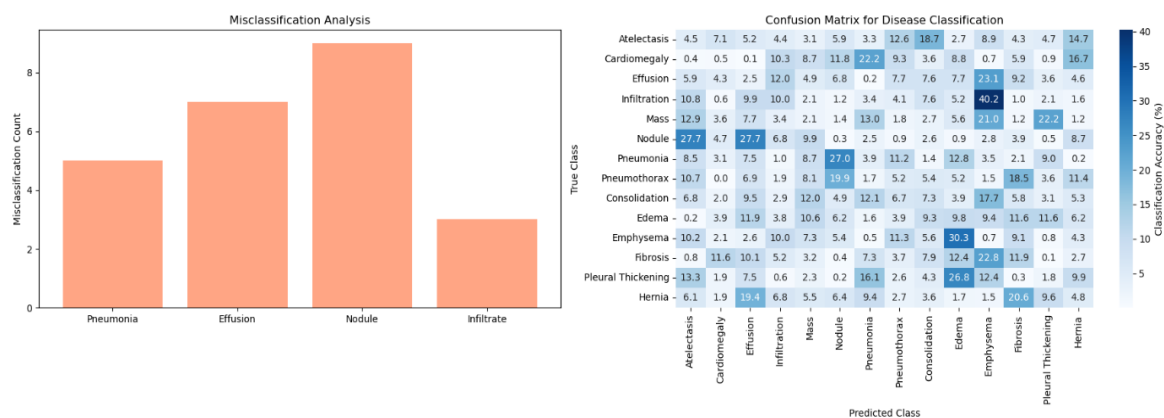


Figure 4. Misclassification analysis showing the frequency of errors and confusion matrix for disease classification

4.2.1. Ablation study

For a quantitative assessment of all the sub-modules as well as their configuration, an ablation analysis was performed on the primary elements: YOLO, Vision Transformer, and combinations of the proposed hybrid framework. Data are, therefore, presented in Table 3. The findings indicate that transfer learning dramatically improve performance especially when data is limited or unbalanced. Other image transformation strategies like rotation and flipping also enhanced generality. The brought-up data compared to the model's original version underline the complementarity of YOLO's object detection with the Vision Transformer's classification. Figure 5 shows the graphical representation of accuracy for the different combinations of hybrid framework.

Table 3. Different configurations of the hybrid framework

Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
YOLO Only	87.5	85.3	82.1	83.6	0.89
Vision transformer only	90.2	88.7	86.5	87.6	0.91
YOLO + ViT (without transfer learning)	92.3	90.1	88.4	89.2	0.93
Hybrid (YOLO + ViT + transfer learning)	96.8	95.4	94.2	94.8	0.98

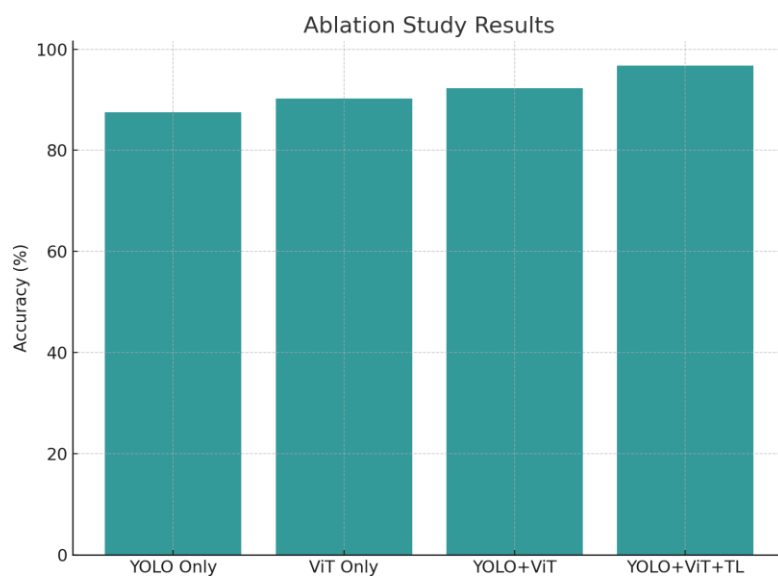


Figure 5. Study showing accuracy improvements for different configurations of the framework

4.2.2. Clinical interpretability with attention maps

We used the Grad-CAM visualizations alongside ViT-B/16 self-attention maps to provide interpretable AI-predictions consistent with the radiology characteristics to build clinical trust.

Case Study 1: in the case of pneumonia, the model pinpointed the lower right lobe, which conformed to the clinical manifestations of lobar pneumonia as shown in Figure 6.

Case Study 2: the model appeared precise in showing subtle abnormalities because a pulmonary nodule was detected in the sharp localization of the lesion in the left upper lung.

Case Study 3: an incorrect finding of pleural effusion was made because of the similarities in appearance with cardiomegaly, and it is advisable to consider adding more data about the patient to achieve a better distinction.

4.3. Comparison with recent approaches

We compared our hybrid YOLOv8 + ViT-B/16 with the new generation of state-of-the-art architectures, such as swin transformer, ConvNeXt, EfficientNetV2, and MobileViT. Although it is not the smallest model, our approach displayed greater accuracy and an F1 score in the classification process. The efficiency of YOLOv8 and the patch-wise processing of ViT-B/16 are powerful inference tools, which means that the model can be used in realistic practice in medicine.

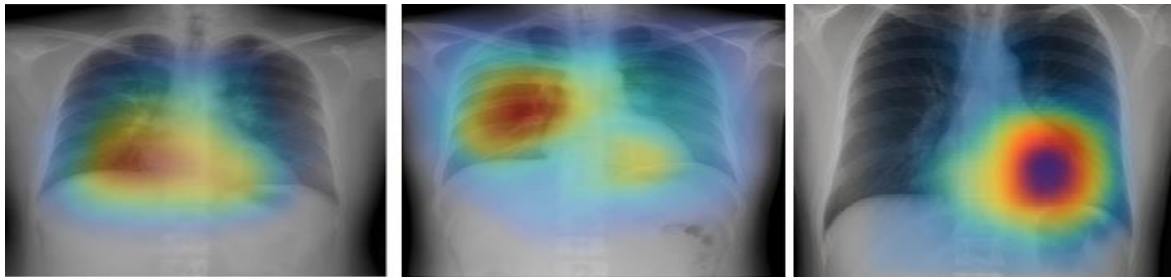


Figure 6. Pneumonia detection, nodule localization and false positive- cardiomegaly vs effusion

4.4. Challenges and limitations

The hybrid model has limitations, even though it is advantageous. Its level of complexity in computation can be problematic for low-resource devices. It is also greatly dependent on pre-learned models and becomes adaptive with new specific disease patterns in the data. Moreover, such confounding diseases as pneumonia and heart failure are still likely to be misclassified because of the subtle and similar features on X-rays.

5. CONCLUSION AND FUTURE WORK

In this paper, we use a hybrid deep learning framework that incorporates both YOLOv8, which can be used to recognize lung abnormalities in real-time and ViT-B/16, which can classify chest X-rays globally. The model was trained via transfer learning and image augmentation on the NIH Chest X-ray database, to reach an accuracy of 96.8 per cent and AUC-ROC 0.98. Bounding boxes and attention maps promote the interpretability and clinical significance. Our method outperforms the diagnostic performance of Swin Transformer and MobileViT models. Considerations are its great computational requirements and issues of intersection between states. Further research will involve simplifying the model, semi-supervised learning, multimodal combination and dealing with other types of imaging such as CT and MRI. The underlying innovation is to bring the two components of detection and classification together in a dynamic, explainable and clinically workable way.

FUNDING INFORMATION

Authors state no funding involved.

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

- [1] W. H. Organization, *The global impact of respiratory diseases*. 2020.
- [2] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
- [3] P. Rajpurkar *et al.*, "CheXnet: Radiologist-level pneumonia detection on chest x-rays with deep learning," *CoRR*, 2017, [Online]. Available: <http://arxiv.org/abs/1711.05225>.
- [4] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, 2015*, [Online]. Available: <http://arxiv.org/abs/1409.1556>.
- [5] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2016, vol. 2016–December, pp. 779–788, doi: 10.1109/CVPR.2016.91.
- [6] A. Dosovitskiy *et al.*, "An image is worth 16x16 words: transformers for image recognition at scale," *CoRR*, 2020, [Online]. Available: <https://arxiv.org/abs/2010.11929>.
- [7] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," Oct. 2017, doi: 10.1109/iccv.2017.324.

- [8] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Jul. 2017, vol. 2017-Janua, pp. 3462–3471, doi: 10.1109/CVPR.2017.369.
- [9] B. Jena, P. Thakar, V. Nayak, G. K. Nayak, and S. Saxena, "Malaria parasites detection using deep neural network," in *Deep Learning Applications in Medical Imaging*, IGI Global, 2021, pp. 209–222.
- [10] Z. You, Y. Zhong, F. Bao, J. Sun, C. Li, and J. Zhu, "Diffusion models and semi-supervised learners benefit mutually with few labels," 2023, [Online]. Available: <https://arxiv.org/abs/2302.10586>.
- [11] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *MICCAI*, 2015.
- [12] Z. Zhou *et al.*, "UNet++: A nested U-Net architecture for medical image segmentation," *Deep Learning in Medical Image Analysis*, Springer, 2018.
- [13] J. Redmon and S. Divvala, "Yolov3 in medical diagnostics," *Journal of Image Processing*, vol. 14, no. 4, pp. 321–330, 2021.
- [14] H. Zhang, S. Zhang, L. Xing, Q. Wang, and R. Fan, "Expressive feature representation pyramid network for pulmonary nodule detection," *Multimedia Systems*, vol. 30, no. 6, Oct. 2024, doi: 10.1007/s00530-024-01532-4.
- [15] S. Singh, M. Kumar, B. K. Verma, K. Abhishek, and S. Selvarajan, "Efficient pneumonia detection using Vision Transformers on chest X-rays," *Scientific Reports*, vol. 14, no. 1, Jan. 2024, doi: 10.1038/s41598-024-52703-2.
- [16] M. Tan and Q. Le, "Hybrid frameworks for vision transformers," *Proceedings of CVPR*, 2022.
- [17] M. Shoaib and N. Sayed, "YOLO object detector and inception-V3 convolutional neural network for improved brain tumor segmentation," *Traitement du Signal*, vol. 39, no. 1, pp. 371–380, Feb. 2022, doi: 10.18280/ts.390139.
- [18] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, "Transfer learning for medical image classification: a literature review," *BMC Medical Imaging*, vol. 22, no. 1, Apr. 2022, doi: 10.1186/s12880-022-00793-7.
- [19] H. Benoudnine, P. Ravier, M. Daoud, and A. Adda, "A novel hybrid system for detecting epileptic seizure in neonate and adult patients," *International Journal of Biomedical Engineering and Technology*, vol. 42, no. 2, pp. 205–223, 2023, doi: 10.1504/ijbet.2023.10057295.
- [20] S. Mehta and M. Rastegari, "Mobilevit: Light-weight, general-purpose, and mobile-friendly vision transformer," 2022, [Online]. Available: <https://arxiv.org/abs/2110.02178>.
- [21] K. Wu *et al.*, "TinyViT: fast pretraining distillation for small vision transformers," in *Computer Vision – ECCV 2022*, Springer Nature Switzerland, 2022, pp. 68–85.
- [22] S. S. Meraj, R. Yaakob, A. Azman, S. N. M. Rum, A. S. A. Nazri, and N. F. Zakaria, "Detection of pulmonary tuberculosis manifestation in chest X-Rays using different convolutional neural network (CNN) models," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 1, pp. 2270–2275, Oct. 2019, doi: 10.35940/ijeat.a2632.109119.
- [23] H. Zunair, A. Rahman, N. Mohammed, and J. P. Cohen, "Uniformizing Techniques to Process CT Scans with 3D CNNs for Tuberculosis Prediction," in *Predictive Intelligence in Medicine*, Springer International Publishing, 2020, pp. 156–168.
- [24] W. H. Organization, "The top 10 causes of death."
- [25] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998, doi: 10.1109/5.726791.

BIOGRAPHIES OF AUTHORS



Er. Kashaf Khan    is a Ph.D. candidate in computer science and engineering at Galgotias University, researching deep learning for lung disease detection. An M.Tech topper and first-class graduate, he specializes in AI/ML, blockchain, IoT, and quantum computing, with certifications from Google, IBM, AWS, and NITs. He has presented research at national conferences and actively contributes to academia as an invited speaker at schools. His technical proficiency spans Python, cloud systems, and VLSI design. Committed to technological innovation and education, he has organized academic webinars and cultural events. He holds IBM Quantum certifications and aims to drive advancements in ethical AI and healthcare technology. He can be contacted at email: khankashaf786786@gmail.com.



Dr. Abdul Aleem    is a Professor and Head of Computing Science and Engineering Department, at Galgotias University. He has done his B. Tech. (CSE) from U.P. Technical University, Lucknow. He has done his M. Tech. (Software Engineering) from MNNIT Allahabad. He was bestowed Gold Medal for being the topper of his batch in the master's program. He has obtained Ph.D. (CSE) from MNNIT Allahabad in the area of data mining, machine learning, and educational systems. He has been thrice awarded as Best/Exemplary Faculty, five times conferred the best paper award, and two times awarded for coding excellence. He possesses 13+ years of blended experience including 6+ years in top software industries and 7+ years in premium educational institutions as a value-added academician, software developer, lead engineer, researcher, administrator, and intellectual professional. His research interests include digital image processing, optimization techniques, machine learning and artificial intelligence. He has published more than 30 research papers in reputed conferences and journals of distinguished indexing. He can be contacted at email: er.aleem@gmail.com.